



THATCamp Paris 2012 Non-actes de la non-conférence des humanités numériques

Éditions de la Maison des sciences de l'homme

Les *digital humanities* sont-elles solubles dans le Steampunk ?

Conférence inaugurale

Paul Bertrand

DOI : 10.4000/books.editionsmsmh.327
Éditeur : Éditions de la Maison des sciences de l'homme
Lieu d'édition : Paris
Année d'édition : 2012
Date de mise en ligne : 1 octobre 2012
Collection : La Non-Collection
ISBN électronique : 9782735115273



<http://books.openedition.org>

Référence électronique

BERTRAND, Paul. *Les digital humanities sont-elles solubles dans le Steampunk ? : Conférence inaugurale*
In : *THATCamp Paris 2012 : Non-actes de la non-conférence des humanités numériques* [en ligne]. Paris : Éditions de la Maison des sciences de l'homme, 2012 (généré le 01 mai 2019). Disponible sur Internet : <http://books.openedition.org/editionsmsmh/327>. ISBN : 9782735115273. DOI : 10.4000/books.editionsmsmh.327.

Ce document a été généré automatiquement le 1 mai 2019.

Les *digital humanities* sont-elles solubles dans le Steampunk ?

Conférence inaugurale

Paul Bertrand

NOTE DE L'ÉDITEUR

Transcription de la conférence inaugurale. Celle-ci peut être écoutée via un lecteur audio : <http://archive.org/embed/LesDigitalHumanitiesSont-ellesSolublesDansLeSteampunk>

- ¹ Le plus important pour moi d'abord ici, c'est de vous remercier et de remercier les organisateurs. Je remercie Pierre en particulier pour son invitation qui m'honore beaucoup et qui m'a fait me poser beaucoup de questions. Dans mon expérience des tous premiers THATCamp, je n'avais pas grand souvenir de conférences inaugurales. Je n'ai pas participé à celle de 2010 ici, mais pour ce que je connaissais, cela ne me disait rien. Je me suis donc dit deux choses : la première, soit cela ne m'avait fait absolument aucun effet, et donc je sais bien ce que je ne dois pas faire, ou plutôt je vais essayer de faire autre chose. Et la seconde, c'est que cela soit original ou pas, c'est un défi. Faisons donc une non-conférence inaugurale ! Ce qui somme toute, est, me connaissant, relativement faisable.

Paul Bertrand



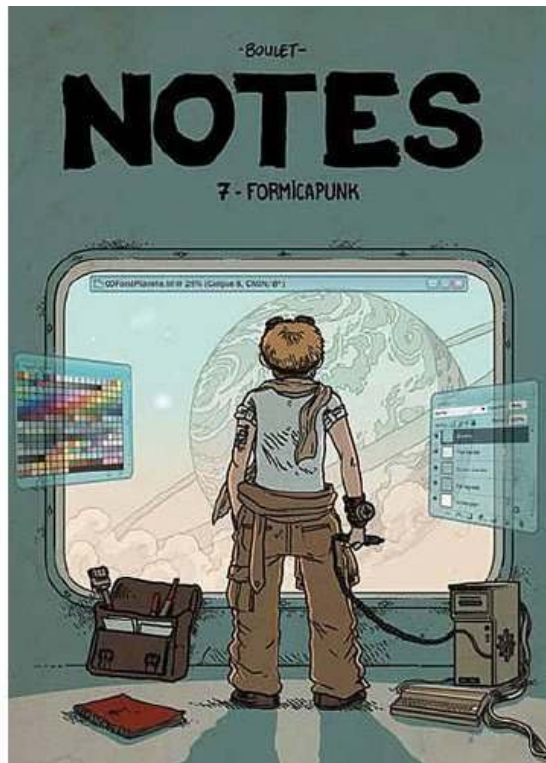
- 2 Dans la mesure où il faut partir d'un concept, le concept que j'ai choisi est celui du « *steampunk* ». Cela, pour deux raisons : la première pour rendre hommage à l'un de mes collègues de la blogosphère avec lequel nous avons commencé, il y a bien longtemps, et qui est le dessinateur Boulet dont le septième tome des *Notes* vient de sortir. Boulet est un grand adepte de cette notion. Pour ceux qui connaissent un peu le monde du blog, c'est un personnage ; et ici je lui rends hommage, vous pouvez le twitter, désormais je suis prêt à devenir son ami et à aller boire des verres avec lui ! La seconde chose, la plus importante, est que le *steampunk* est un mode de réflexion, plutôt une esthétique, très intéressante qui est en train de prendre de l'importance ces dernières années – depuis déjà au moins une petite vingtaine d'années, mais plus fortement ces dernières années. Au départ, c'est un genre littéraire, un sous-genre de la science-fiction dite « uchronique » – c'est une définition que j'ai récupérée à partir de Wikipédia, c'est pour vous dire si c'est sérieux ! – dont l'intitulé a été forgé par allusion au « cyberpunk ». Pour cette raison, il est plus approprié de parler de « rétro-futurisme ». L'expression *steampunk*, je continue la définition, est un terme qui signifie « punk à vapeur », au cas où vous auriez un souci de traduction, et peut être traduite par « futur à vapeur », terme inventé pour qualifier un genre de la littérature de science-fiction, né à la fin du XX^e siècle, dont l'action se déroule dans l'atmosphère de la société industrielle du XIX^e siècle. Le terme fait référence à l'utilisation massive, je cite toujours, des machines à vapeur au début de la révolution industrielle et à l'époque victorienne.

aurelberra

RT @jpmasse : @openeditionsays en latin ?

c'est totalement *steampunk* non ? ;-) #tcp2012

Notes 7 : Formicapunk, de Boulet



- 3 À l'origine, c'est un genre littéraire. Mais le plus intéressant, c'est de voir que très rapidement c'est devenu un genre en soi qui s'est répandu dans la société, et est devenu un mode de représentation esthétique. Vous savez que de nos jours, l'uchronie a beaucoup de succès, et parmi les formes d'uchronies, le *steampunk* en est une très importante. Une des différences principales entre le *steampunk* et la science-fiction c'est la présence d'éléments anachroniques plus tardifs, comme les ordinateurs, les manipulations génétiques, ce qui permet d'y développer tout un humour spécifique... L'idée est donc de créer un décalage ; un décalage entre, d'une part un contexte dix-neuvième siècle, et de l'autre toute une technologie, ou toute une aspiration à un progrès technologique qui lelle est plutôt avant-gardiste. Pour ceux qui ont quelques connaissances de la série *Star Trek* à la télévision, il y a des épisodes qui sont déjà clairement *steampunk*, où l'on voit des gens vivant façon dix-neuvième siècle (avec l'œil américain évidemment), mais qui à côté de cela arrivent à développer des armes de destruction massive avec pratiquement deux bouts de caoutchouc ou des bouts de bois. Il est donc très intéressant de voir que ce système esthétique est en train de se développer.
- 4 Ce qui est encore plus intéressant, c'est que ce style se rapproche très fortement, par certains côtés esthétiques, de la vision actuelle que l'on peut avoir du numérique et de son évolution. Je vais vous citer une partie de la définition que j'avais trouvée dans Wikipédia : « tandis que l'ère de l'électronique tend vers la miniaturisation – grand fantasme – l'élément esthétique fondamental du *steampunk* est son gigantisme » : demandez à l'IN2P3 et sa salle des serveurs ! « Cette technologie, délibérément bloquée au stade de la vapeur, produit des constructions démesurées, complexes, extravagantes de tuyauteries actionnées par des leviers et des claviers aussi nombreux que compliqués. Les véhicules à vapeur semblent aussi lourds et difficilement maniables, ils nécessitent une très forte dépense d'énergie pour un rendement faible. Et surtout, enfin, le héros

classique dans ce genre d'univers est le mécanicien de génie », on aurait pu dire le *geek*, cela aurait été aussi bien.

- 5 À côté de cela, continuons dans les définitions, les *digital humanities* désignent « une transdiscipline porteuse de méthodes, de dispositifs et de perspectives heuristiques liées au numérique dans le domaine des sciences humaines et sociales. [...] Nous appelons à l'intégration de la culture numérique dans la définition de la culture générale du XIX^e siècle. » Vous avez reconnu le Manifeste des digital humanities !
- 6 Je voulais poser ces deux éléments de définition, parce que maintenant je vais les faire se télescoper. Ce dont je vais parler ici, c'est de ma vision, totalement subjective qui est le résultat de mon expérience et de mes connaissances (quelles que soient leurs qualités) des *digital humanities* depuis maintenant presque une dizaine d'années, et même davantage si remonte aux débuts quand je commençais à bricoler avec des bases Access. Je vais donc mobiliser cette expérience, en me référant évidemment au *steampunk*, mais pas seulement, pour évoquer ma vision de l'évolution actuelle des *digital humanities*. C'est une vision qui paraîtra par certains aspects dure voire négative, mais je vais essayer de présenter les choses telles que je les ressens. L'objectif c'est de vous faire bouger, de vous faire réagir.

Première couche : Une insoutenable tension dialectique : autour d'une esthétique du *steampunk*

- 7 La première couche d'interprétation du *steampunk*, forcément, c'est celle qu'avait donné en son temps dans un congrès sur les Linked Open Data for Ancient Studies¹, Andrew Reinhard. Je ne suis donc pas le premier à avoir parlé de *steampunk* pour les *digital humanities*. Andrew Reinhard avait lancé un pavé dans la mare en avançant l'idée suivante : « [...] the current scholarly publishing enterprise is essentially *steampunk*, 21st century work with 19th century models [...] » Il y a une certaine violence dans cette déclaration ; et en même temps un certain réalisme.
- 8 Si je me réfère à mon expérience, les modèles éditoriaux que je retrouve en règle générale dans les équipes avec lesquelles j'ai travaillé jusqu'ici correspondent à l'état d'esprit décrit par Andrew Reinhard - avant de me retrouver à l'université de Louvain, j'ai été comme certains d'entre vous le savent peut-être, et je le suis toujours d'ailleurs en grande partie, rattaché à l'Institut de recherche et d'histoire des textes (IRHT) ; j'ai dirigé, et je dirige toujours, la partie scientifique du centre de ressources numériques Telma. On arrive avec une édition, souvent d'une grande qualité scientifique, que l'on veut à tout prix publier sur le web, mais en PDF. Et on la met en ligne en l'état parce que l'on n'a pas le choix, sous l'effet d'une pression qui vient de tous les niveaux de la recherche, et même si on peut avoir d'autres perspectives et d'autres points de vue sur cette publication.
- 9 Je ne donnerai aucun exemple particulier dans les quelques mots qui vont suivre parce que, vous me connaissez, j'ai horreur de dire du mal des gens ! Je me contenterai d'un commentaire long sur quelques éléments que j'ai pu retirer ici ou là, autour de points de changement nécessaires dans cette vision des *digital humanities* qui se développe actuellement.
- 10 Le premier point, est de savoir comment régler cette question de l'urgence de la publication, et ce divorce constant entre chercheur et ingénieur.

NuevoMundoMN

El divorcio entre el investigador y el ingeniero :

¿un punto clave de las humanidades digitales ? #tcp2012

- 11 Il y a pour moi une tension dialectique qui est perpétuée, que l'on ne retrouve pas seulement dans la publication d'un corpus quel qu'il soit. Je vous ai parlé des publications en PDF, mais je pourrais évoquer aussi des corpus publiés dans des formats encore plus frustrés ou encore plus problématiques, soit parce qu'ils sont trop anciens, soit parce qu'ils sont structurés selon des modèles qui nous font hérisser le poil... jusqu'à voir des éditions publiées sur le web en .doc au prétexte que la fonction CTRL-F est largement suffisante ! Et cette tension se reproduit également dans d'autres domaines, comme par exemple les bibliothèques.
- 12 Il est vrai que les bibliothèques ont dix longueurs d'avance par rapport au monde de la recherche, du point de vue de l'appropriation des nouvelles techniques du web sémantique, de l'appropriation des concepts de bonnes pratiques ou d'archivage pérenne, ou encore de l'utilisation des moteurs ou des ontologies. N'empêche que beaucoup de bibliothèques, surtout à des niveaux locaux, car je ne parle pas de machines de guerre comme la BnF, subissent une pression qui vient de leur tutelle, par les collectivités elles-mêmes, pour que les collections soient publiées de manière très rapide, très frustrée, éventuellement dans des formats qui sont comparables à ceux dont je parlais tout à l'heure, avec des métadonnées frustrées également : le Dublin Core du bas du tableau. Quand certaines bibliothèques ou dépôts d'archives ne disent pas tout simplement : « une partie seulement est visible sur le web, et pour le reste venez chez nous pour avoir accès à tout ». Donc, une sorte de libre accès à deux vitesses qui se trouve malheureusement être une constante, et qui est non pas commandé par les bibliothèques elles-mêmes ou les archives mais par les collectivités. Il y a donc un problème entre une demande d'un côté et, d'autre part, une capacité de production, des règles, un règlement, une organisation, qui eux se trouvent en porte-à-faux.

2. De grands espoirs : la dissémination et l'appropriation

- 13 Nous avons de grands espoirs pour régler le problème de l'utilisation de formats nouveaux, pour mettre en œuvre ces fameuses bonnes pratiques qu'on connaît bien maintenant, pour mettre en œuvre tout une série de procédures d'archivage pérenne, pour participer aux grandes procédures d'interopérabilité : on avait beaucoup d'espoirs que cela puisse décoller assez vite.

THATCamp Paris 2012



- 14 Voilà maintenant à peu près cinq années que l'on peut dire que les *digital humanities* ont pignon sur rue en France. Plus, si on compte les pratiques sur le terrain. Là encore, je prends l'exemple de ce que je connais bien : je me souviendrai je crois toute ma vie du moment où nous avons lancé la plate-forme Telma. Pendant des années cette plate-forme a peiné à démarrer parce que l'on n'avait pas de reconnaissance institutionnelle. C'était un partenariat École des Chartes / IRHT qui fonctionnait très bien du côté de l'École des Chartes tandis que de celui de l'IRHT, on ne comprenait pas très bien à quoi cela pouvait servir. Et c'est lorsque nous avons réussi à faire une démonstration, deux ans après la première publication, que l'on a réussi à convaincre *in fine* la direction (à partir du moment où elle partait, c'est cela qui est le plus intéressant !) de l'importance de l'objet. Parce qu'en parallèle se mettait en place toute une série d'autres projets similaires un peu partout en France, la direction se rendait compte de l'importance de l'initiative. Et tout à coup, toutes les portes nous ont été ouvertes. À tel point que la grosse verrue sur le dos de l'IRHT, si je puis dire, est devenue, tout d'un coup, un fleuron ; une arme importante pour la grandeur de l'institution.

Les *digital natives*

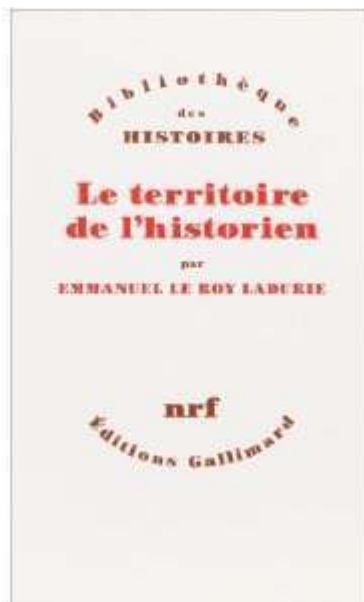
- 15 Il y a quatre ou cinq ans, on misait beaucoup sur les *digital natives*, ces jeunes qui sont nés dans le monde numérique et que l'on retrouve maintenant à l'université, que ce soit à Louvain ou ailleurs. Pour ma part, je suis revenu des *digital natives* qui sont plutôt des *Facebook natives*. De fait, on est très loin de l'appropriation du numérique en tant que tel. Lorsque l'on propose à des chercheurs, très jeunes, qui viennent d'arriver, qu'ils soient en thèse ou juste après, de se lancer, ils peuvent ouvrir un carnet de recherche – ce qui correspond bien en quelque sorte aux pratiques de réseau social qui sont en train de se

développer et c'est une excellente chose. Mais lorsqu'il s'agit d'aller plus loin dans l'appropriation des *digital humanities*, même de manière assez distante, il y a comme un blocage et une interrogation. Et l'appropriation est lente et difficile.

Les historiens programmeurs de Le Roy Ladurie ?

- 16 Deuxième espoir que nous avons, les historiens programmeurs, et je sais que cela a fait l'objet de beaucoup de discussions ces derniers temps. Sur ce concept aussi, je suis très perplexe. Je jette ma pierre dans le marigot de la discussion dans laquelle tout le monde se dépêtrera demain. Je pense que ce concept d'historien programmeur est dangereux et n'a pour l'instant pas donné de résultats. En tous les cas, comme le concevait Le Roy Ladurie, c'est-à-dire comme un historien tout à coup capable de jongler, qui avec du Python, qui avec du Java, voire qui avec du Perl. Car le problème est que beaucoup de ces historiens programmeurs, - on en connaît actuellement qui ont en effet appris la programmation - programment fort bien, mais qu'un historien a une vie d'historien déjà bien remplie en soi, une vie de recherche. Lorsqu'à un moment ou un autre il se lance en profondeur dans une démarche d'apprentissage technologique (un langage, une méthode de programmation), il s'y lance à fond, et applique cette technique à son objet de recherche. Mais, vient alors le reste de la vie. Et le reste de la vie, ce sont les charges d'enseignements, la recherche qui se développe, les projets d'ANR. Et il n'a plus la possibilité matérielle de développer ce savoir, de le faire évoluer, de mettre en place pour lui-même une veille et continuer à développer ce savoir faire technologique. Ce savoir se retrouve donc bloqué à un instant *t*.

Le Territoire de l'historien



- 17 Cela ne serait pas grave si l'historien lui-même acceptait l'idée que ses connaissances technologiques peuvent être dépassées. Or le problème est que beaucoup de ces historiens, pour l'instant en tous cas, sont persuadés que le langage qu'ils connaissent et qu'ils maîtrisent parfaitement est le « couteau suisse » parfait. Il y a ainsi aujourd'hui des chercheurs qui bloquent des processus d'avancement dans une discipline parce que l'on ne peut pas arriver à les faire passer de cet outil à autre chose. J'ai un exemple récent

concernant le passage d'une base de données codicologique et paléographique sur un autre support PHP : « Oui mais tout de même, j'aimais bien mon ancien système. Ne peut-on pas le laisser à côté, pour tous les autres chercheurs qui y sont habitués ? » Il y a une difficulté terrible à évoluer et à passer à autre chose. Le risque est donc un blocage dans cette évolution, ce processus de construction des *digital humanities*, avec des chercheurs qui ne voudront pas tout simplement évoluer vers de nouvelles étapes de développement.

- 18 C'est la raison pour laquelle pour moi la notion d'historien programmeur reste dangereuse. Où alors il faut envisager un historien avec une teinture de *digital humanities* assez forte pour en comprendre les enjeux, et qui suive cette teinture en étant en contact direct avec des gens qui puissent lui instiller régulièrement, lui faire prendre conscience de l'évolution des choses et lui donner les moyens d'une veille rapide et efficace. En fait, ce dont les chercheurs ont besoin, c'est de se dépêtrer dans la masse de d'informations produites autour des *digital humanities*.

piotrr70

Besoin énorme pour les chercheurs de se repérer
dans le foisonnement des DH #tcp2012

spouyllau

@piotrr70 il faut des boussoles en fait #tcp2012

- 19 Je suis ravi du nombre de colloques, du nombre de conférences, du nombre de journées d'étude, du nombre de symposium qui s'organisent. C'est parfait, c'est la preuve que quelque chose se passe. Mais un chercheur lambda, quel qu'il soit, ne peut pas se débrouiller dans tout cela. Comment voulez-vous qu'il arrive à trouver les quelques éléments qui puissent lui montrer vers quoi on va et ce qu'il faudrait faire ? Il est donc nécessaire de rationaliser l'information.

3. Deuxième couche : qu'est-ce qui coince ? L'idéologie du Steampunk

Un risque d'appauvrissement des données

- 20 Continuons avec autre chose. J'en remets une couche. Jusqu'à présent, on en est resté à un niveau esthétique. Il n'y a pas de problèmes techniques, on n'est pas confrontés réellement à un danger pour la recherche elle-même, mais plutôt à un ralentissement dû à une absence de dialogue. Le problème est que le *steampunk* est aussi une idéologie qui implique une façon de voir, ou de conceptualiser les données historiques, à les figer à un certain niveau. Je pense, vous allez comprendre où je veux en venir, qu'on est arrivé à un certain palier dans le développement des *digital humanities* et que grandit le risque d'un appauvrissement des données.
- 21 On court le risque d'un appauvrissement des données plutôt qu'à un enrichissement de ces données, au moment du passage au numérique. L'appauvrissement, c'est un mot très dur que je vais tenter d'expliquer.
- 22 En fait, je vais tout vous dire : pour préparer cette conférence, je me suis demandé quelle démarche heuristique mettre en oeuvre. Cela serait intéressant aussi pour le THATCamp de savoir comment mettre en place une heuristique *digital humanities* à partir des

données. Comment ai-je préparé ma conférence ? J'ai décidé de prendre trois semaines de tweets. J'ai pris tous les tweets que j'ai reçus depuis trois semaines et j'ai été voir chacun des liens vers lesquels on me renvoyait en me disant que cela me faisait un état, qui vaut ce qu'il vaut, mais un état de la situation de la réflexion de tous ceux qui travaillent dans les *digital humanities* que je connais et avec lesquels je travaille. Un certain nombre d'éléments m'ont sauté aux yeux et je vous les livre ici.

Vive le texte !

- 23 Le premier élément concerne la notion de « texte ». Je pense en particulier à une façon de voir les choses qui a été développée par un chercheur anglo-saxon du nom de Tim Hitchcock sur son blog (c'est un lien qui je crois m'avait été renvoyé par Frédéric Clavert). Cette note de blog disait clairement que l'on n'avait plus besoin du livre. En effet, lorsque le texte passe dans le monde des *digital humanities*, le concept de livre devient secondaire et on entre dans une autre perspective fondée sur les données. Cela signifie que le métier d'historien lui-même qui était basé sur le livre et son appropriation par le chercheur – je parle de l'historien mais on pourrait évoquer n'importe quel chercheur en sciences humaines –, le chercheur qui s'appuie sur concept fondamental du livre comme élément de référence, voit son métier déconstruit en quelque sorte par la disparition d'un objet emblématique pour lui, progressivement remplacé par le texte.
- 24 Le problème est que le texte n'existe pas. Il n'a jamais existé par lui-même. Le texte n'existe que parce qu'il a été conçu dans un cadre précis ; un livre, un article, une note de blog éventuellement. Mais le texte est conçu de manière complètement liée intellectuellement à l'objet matériel de départ, physique ou numérique. Donc, le livre n'est en quelque sorte pas mort, si l'on prend en considération l'objet de départ. Un texte est lié à son contexte culturel de lecture, de production, et on ne peut pas le dégager de ce contexte. Le risque que nous prenons tous avec les *digital humanities* telles qu'on les voit se développer aujourd'hui est que le texte soit détaché de son contexte de production.

Neuromancien

paul bertrand considère que le texte n'existe pas
indépendamment et qu'il est lié à l'objet matériel
de départ #tcp2012

Cclivaz

#tcp2012 Fahrenheit 451 est un roman paradigmatique
de la culture digitale : c'est l'humain qui porte la culture,
quels que soient les médias

Vive la donnée !

- 25 L'autre élément sur lequel je voudrais revenir est la donnée. Là encore, on constate des fantasmes. Tout le monde nous parle des *data*. La donnée est devenue un objet absolument essentiel, on est certain qu'elle est devenue l'élément premier de notre recherche, et qu'avec cette donnée nous allons pouvoir, en l'organisant, en l'interrogeant, en l'enrichissant par des procédures du web sémantique, produire des résultats, des idées, de la nouveauté intellectuelle.

- 26 Or le problème est que la donnée n'existe pas. La donnée cela n'existe pas ! De quoi s'agit-il ? Est-ce un texte, une ligne de texte, un mot, un concept ? Est-ce que c'est un concept derrière un concept, est-ce que c'est une idée derrière un concept ? Je sais bien, je ratiocine, mais tout de même, je demande qu'on me donne une explication claire de ce qu'est une donnée.

Merryarchive

#tcp2012 P. Bertrand. Oh là là, la donnée n'existe pas non plus !

Inactinique

Lisez, sur les données, l'article de danah boyd

dans le ReadWriteBook 2 #tcp2012

- 27 Une base de données est un outil extrêmement utile, mais elle crée une information qui elle-même n'existait pas jusque-là. Prenons l'exemple d'une base de données autour de la codicologie. On part de deux manuscrits en tant que tels et on commence à prendre des informations codicologiques autour de chaque manuscrit ; on en fait une très belle base codicologique sur les reliures, ou bien sur les réglures ; pourquoi pas, chacun ses fantasmes ! Le problème est que la donnée codicologique est une construction pure, elle est récupérée de l'objet livre dans des conditions qu'il faudrait encore pouvoir définir, et l'objet livre n'est plus lui-même dans la donnée. Cela signifie que l'on va interpréter les résultats de recherches statistiques en les détachant de la réalité du livre lui-même, de la réalité du manuscrit. Et forcément, on ne pourra pas parler d'une donnée complète qui pointerait complètement vers toutes les réalités du manuscrit.
- 28 Or, la nouveauté en recherche, on le sait bien, ne vient pas de la mise en relation de données dont on se doute bien qu'en les mettant ensemble on arrive à des résultats, mais de la mise en relation de données à propos desquelles on n'avait aucune idée de leur possible mise en relation. Le risque, avec la création de cette donnée, c'est de n'avoir qu'une donnée partielle et souvent non définie. Qu'est-ce qu'elle représente ? Est-ce que c'est une source, une partie de la source, est-ce que c'est une métadonnée au-dessus de la source ? Lorsque l'on étudie un objet documentaire, on peut le découper en autant de données que notre imagination le permet et ce concept de données n'est pas posé de manière complète.

Vive les *big data* !

- 29 L'autre grand fantasme, ce sont les *big data*, les grandes masses de données qui font évidemment beaucoup rêver de l'autre côté de l'Atlantique mais qui ici aussi commencent. C'est tout le problème de l'histoire du chiffre. En sciences humaines, il y a 30 ans, on ne jurait que par la statistique et par le quantitatif. On a depuis jeté aux orties le quantitatif de manière très virulente, et via les *big data*, il semble revenir par la fenêtre ; Pourquoi pas après tout ? Le problème est que ces énormes masses documentaires aplatissent la qualité des données, ou en tous cas l'information de base. La critique a d'ailleurs été bien posée dans l'article publiée dans le Read/Write Book 2 : les *big data* sont à la fois incontournable et en même temps très dangereuses². Il n'y a rien de plus difficile à analyser, pour moi, qu'une série de chiffres.

EmmaCorne

il y a 20 ans en shs on ne jurait que
par le quantitatif, on l'a viré et là il revient
par la fenêtre @PaulBertrand #tcp2012

Vive les plate-formes et les méta-moteurs !

- 30 Evoquons les métadonnées, très rapidement. Parce que là aussi c'est encore un fantasme. Les métadonnées sont souvent la seule clef d'accès, à l'heure actuelle, pour bien des documents et bien des textes. Le problème est que les métadonnées que l'on utilise actuellement sont souvent très faibles. La plupart du temps, et je parle d'expérience, j'ai vu les métadonnées qui sont utilisées dans les disciplines où je travaille, et il faut bien avouer que c'est un peu triste. Il est vrai qu'heureusement les plate-formes et les méta-moteurs existent, et on ne peut qu'en être satisfait, tout en même temps en restant extrêmement prudent car ils dépendent en grande partie de leur qualité de contenu, de la qualité du marquage. C'est vrai – et là je m'incline devant Stéphane Pouyllau qui est là – qu'Isidore³ résout déjà une partie du problème grâce à son enrichissement sémantique. Il n'empêche que – et là, je crois qu'il le reconnaîtra avec moi – la qualité d'Isidore reste encore très variable. Oui évidemment, je suis persuadé qu'on peut l'utiliser pour la recherche. Mais comment l'utiliser ? Et comment améliorer les métadonnées que ces plate-formes et méta-moteurs utilisent ?

Vive les ontologies !

- 31 On a enfin les ontologies, un vieux problème en discussion depuis de nombreuses années. Parce que hélas, elles sont extrêmement complexes et extrêmement lourdes à mettre en place, souvent au cas par cas. Moi je n'ai pour l'instant pas encore vu beaucoup d'ontologies fonctionner de manière ample dans nos disciplines. Je ne dis pas qu'ailleurs il n'y en a pas. Mais les quelques rares ingénieurs de *digital humanities* qui s'y collent me regardent souvent avec des yeux effarés quand je leur demande : « Alors, quand est-ce que l'on va pouvoir travailler avec vos ontologies, quand est-ce que vous nous aidez à les mettre en place ? »
- 32 Lorsqu'on parle d'interopérabilité par les ontologies, je me demande si on ne se paie pas de mots. Le risque que j'ai vu apparaître dans des laboratoires très proches, où l'on me dit qu'il faut développer l'interopérabilité – et je suis le premier à dire, bien sûr c'est une excellente idée – est qu'elle se résume à permettre un accès autour des noms d'auteurs, des cotes des manuscrits ou des livres. C'est-à-dire quelque chose que l'on faisait très bien avec des index traditionnels en papier. . On pourrait aller un peu plus loin, mais comment procéder ? La pression institutionnelle fait que le retour aux index traditionnels pointe à l'horizon.

piotrr70

P. Bertrand ne se paie pas de mots :
il rend compte sincèrement d'interrogations
réelles que se pose un chercheur bien au courant #tcp2012

Vive les bonnes pratiques !

- 33 Enfin, et là je vais jeter un gros pavé dans la mare, il faut parler des bonnes pratiques. Parce que les bonnes pratiques sont peut-être une excellente chose, mais elles sont aussi un frein. Tout d'abord il faudrait voir exactement ce qu'il y a derrière, car c'est aussi un terme dont on se gargarise, il faut bien l'avouer. Par exemple, on se focalise sur l'archivage – je ne dis pas que ce n'est pas important –, on se focalise sur la structuration et la possibilité de mise en opérabilité, et on oublie (ou en tous cas on ne se positionne pas) sur le reste. On ne se positionne pas sur l'innovation conceptuelle par exemple, sur l'enrichissement sémantique, sur la granularité des outils d'interrogation, ou simplement sur la décomposition de ces fameuses données. J'ai peur qu'on se contente de bonnes pratiques aujourd'hui dans les *digital humanities* et que l'on renonce à un certain devoir d'audace.

NuevoMundoMN
Humanidades digitales : imposibles
sin »deber de audacia » #tcp2012

3. Deuxième couche : pourquoi ça coince ?

- 34 Je ne voudrais pas être trop pessimiste et chercher des raisons à ces blocages. J'en vois deux.

Le poids de l'institutionnalisation

- 35 La première c'est le poids de l'institution. Nous sommes dans une société où on se satisfait des bonnes pratiques parce que l'on n'a pas le temps de l'innovation. L'institution, les institutions au sens large du terme, pèsent de tout leur poids sur nos disciplines, nos disciplines de recherche et nos disciplines *digital humanities* avec une exigence de résultats rapides et efficaces. Je le vois, lorsque je travaille avec des développeurs pour les différents projets que l'on a mis en place, que cela soit à l'IRHT, à Telma ou ailleurs, à travers les demandes de chercheurs qui sont soumis aux pressions des financeurs et ce poids fini par retomber sur les épaules des développeurs et des spécialistes des *digital humanities*. Il faut aller vite pour publier quelque chose dans les temps. Alors, on fait du PDF et on réfléchit après, ou alors on fait un peu de TEI-Lite, , et puis on publie. Mais c'est de la TEI *light*, et on ne pourra pas aller plus loin car on n'aura pas le temps d'approfondir. Le poids de l'institution est terrible ; il faudrait pouvoir y réfléchir et s'en dégager.

L'accélération du temps

- 36 Deuxième élément, l'accélération du temps. C'est un concept que je prends au sociologue Harmut Rosa qui a publié un superbe ouvrage il y a un an ou deux sur cette question⁴.
- 37 De la même façon que l'accélération du temps a pesé sur notre manière de voir les choses et sur notre inconscient, l'accélération du temps pèse sur nos méthodes de construction de projets, sur les projets eux-mêmes et la nécessité de résultats dont je parlais tout à

l'heure. On peut parfaitement mettre en place un projet de recherche ANR sans jamais avoir jamais rencontré les personnes avec qui on travaille ; je le sais bien, je l'ai fait.

- 38 Donc, comme disait Frédéric Clavert sur Tweeter, on est dans du « post-post-modernisme », on ne reconceptualise plus la donnée parce que l'on n'en a plus le temps et plus les moyens. On revient à une situation critique positiviste, similaire de celle d'avant l'Ecole des Annales, utilisant les documents comme autant de carrières de faits, et figeant ces faits en les atomisant tout simplement. On n'a pas le temps d'aller faire autre chose. On est dans un positivisme si fort, que le chercheur parle encore moins qu'avant avec le développeur ; il n'en a plus le temps.
- 39 En ce sens, les *digital humanities* rejoignent donc non-seulement l'esthétique, mais aussi une idéologie *steampunk* : coincées dans un fatras technologique institutionnel, engoncées dans ces cols amidonnés victoriens où le mécanicien est roi et l'usine à vapeur, pour ne pas dire l'usine à gaz, refont surface. Le risque n'est pas tant de se fondre dans l'idéologie ou l'esthétique *steampunk* – qui est clairement une esthétique de crise, c'est clair – mais c'est surtout de fonder cette idéologie qui prend le pas et qui s'impose, comme naturellement, à beaucoup d'entre nous.

4. Sortir de l'idéologie du *steampunk*

- 40 Alors que faire, car je ne veux pas vous laisser là, tout triste (et surtout je ne veux pas que vous me jetiez trop de tomates quand j'aurai fini de parler). J'ai deux propositions.

jpmasse
#tcp2012 P Bertrand ne veut pas nous laisser tous tristes ;
non à la #tristitude <http://t.co/7ilphrsH>

Reconceptualiser le « donné »

- 41 La première proposition consiste à reconceptualiser le « donné » et pas seulement avec des balises, mais d'un point de vue intellectuel, sur pièces. Il faut une nouvelle critique car la critique sur laquelle nous nous basons date du XIX^e siècle, un moment tout à fait important de l'histoire mais qui n'est plus adaptée à nos besoins. J'ai la chance de donner un cours de critique historique à des étudiants de l'université en Belgique qui est peut-être le seul pays d'Europe où l'on enseigne la critique, en tant que telle, comme discipline. Et je me rends compte que celle que l'on enseigne à l'heure actuelle, suivant les vieux systèmes, Langlois et Seignobos⁵ – enfin on a un peu dépassé, mais pas beaucoup – est complètement à revoir.



- 42 Il faut la revoir de manière ouverte, et en mouvement. Car ce qui est intéressant, c'est de repenser le donné (je dirais, plutôt que les données), à partir de sa situation. Remplacer la

situation de chaque donnée dans son contexte, que ce soit du mot jusqu'au corpus, puis après cela à toutes les déclinaisons du corpus puisqu'une déclinaison du corpus est un nouveau document en soi. Une base de données crée un nouveau donné. Il y a création de quelque chose de neuf.

- 43 Je propose de développer une réflexion un peu similaire à celle que j'avais essayée d'avoir avec Christophe Jacobs, il y a de cela deux ans, pour l'ANR mis en place par Jean-Philippe Jeunet autour du numérique. Une tentative de mouvement critique qui est encore à continuer et à approfondir franchement, notamment parce que le donné ne peut plus seulement être considéré comme un élément que l'on replace dans un contexte, avec un avant, un après, un contenu, un contenant. Mais aussi parce que le donné est en mouvement, en relation. C'est un donné que l'on doit mettre en relation avec d'autres, et soumettre à la critique le type de relation ainsi. Il y a une nouvelle critique documentaire à avec les spécialistes des *digital humanities* qui peut passer par une sémantisation des données, mais qui ne pourra surtout pas faire l'économie d'une reconceptualisation critique.
- 44 Je propose par ailleurs de créer un espace hors de la pression du temps et des institutions, dédié à l'innovation, presque pour le plaisir. Il s'agit de réfléchir avec ceux qui le veulent autour de ce nouveau guide critique du donné tel quel le chercheur en *digital humanities* le voit apparaître et tel que le chercheur lui, de son côté, tente de le mettre en place avec ses propres moyens. C'est un espace de dialogue qui s'établit au niveau de la critique qui peut nous permettre d'échapper à cette idéologie potentielle, et de toute façon certainement de cette esthétique de crise qu'est le *steampunk*.
- 45 Voilà, j'ai beaucoup parlé, maintenant à vous de me jeter les tomates ! Merci.

MajBen
comment faire parler une salle
qui ne f/sait que tweeter #tcp2012

cbarthonnat
Et maintenant, Paul Bertrand (@medieviz)
va analyser tous les tweets relatifs
à sa non-conférence :-D #steampunk #tcp2012

Discussion

- 46 Lou Burnard : Merci beaucoup Paul pour cette intervention très provocatrice. Il me semble que (je ne suis pas historien, donc je me trompe peut-être) mais je crois que la plupart des institutions actuelles, quelques-unes c'est sûr, ont été fondées par le passé en réaction à ce que tu viens juste d'évoquer : un mécontentement avec les institutions du moment. N'échappera-t-on donc jamais à ces cycles de création d'institutions ?
- 47 Paul Bertrand : Oui c'est certain. Jusqu'à un certain point, le THATCamp commence à être une institution.
- 48 Pierre Mounier : Pas pour le moment car les THATCamps sont justement porteurs d'une rhétorique anti-institutionnelle et se définissent un peu comme des zones d'autonomie temporaire.

- 49 Paul Bertrand : Certes, mais c'est une initiative qui a réussi à se positionner dans le temps avec des structures, avec un cadre, avec un logo. Il y a toujours de l'institutionnalisation. C'est pour cette raison qu'il faut créer un espace en dehors du temps qui pourrait même être virtuel, et même s'il y aura une institutionnalisation inévitable avec la production d'un livre ou le partage d'un bon moment. Le plus important pour moi, et c'est seulement un avis, c'est de pouvoir se dégager du poids des exigences d'efficacité et de rapidité de ces institutions qui nous gouvernent, si je puis dire, dans lesquelles nous tentons de travailler et nous tentons d'avancer. Si nous nous débarrassons de cet idée que l'on veut à tout prix avoir vite des résultats, et être très rapidement efficace, alors, on risque de réussir.
- 50 C'est très idéaliste ce que je propose, mais si on n'est pas idéaliste, alors on ne sert plus à rien.
- 51 Jean-Luc Pinol : Je poserai quelques questions pas du tout comme directeur du TGE Adonis, mais comme historien.
- 52 Globalement, je suis historien des villes. La ville est une institution qui se recompose en permanence. Si elle ne se recompose pas elle meurt. C'est un peu la même chose pour les institutions ; il faut qu'elles se reconstituent, se recomposent en permanence. Je ne ferai qu'une petite parenthèse, si l'on regarde le devenir des grandes infrastructures de recherche et la manière dont on repense, au-dessus d'elles, leurs missions, on est en plein dans ces questions.
- 53 La deuxième question est plus étroite et, d'une certaine manière plus corporatiste. Tu as évoqué les chiffres, et tu as dit qu'avec les chiffres on n'a jamais rien démontré. Je suis d'un avis exactement opposé. Précisément, le numérique nous permet avec des méthodes de visualisation de faire des choses totalement différentes et nous oblige à repenser nos données. J'ai vu, comme directeur de laboratoire, des gens qui n'avaient pas quitté le XIX^e siècle. Je suis tout à fait d'accord avec le fait que les étudiants sont bien souvent plus des *facebook natives* que des *digital natives*. La plupart, même s'ils ont 20 ou 25 ans de moins que moi, pensent qu'avec Word on fait tout. Mais c'est vrai qu'il faut réfléchir aux nouvelles manières de penser, en l'occurrence ici, la recherche historique, en utilisant ces nouveaux outils.
- 54 Paul Bertrand : Pour répondre rapidement, sur l'institutionnalisation et sur le fait que l'institution se refait, c'est très bien, et même au contraire c'est une excellente chose. C'est d'ailleurs la raison pour laquelle, la refonte de l'institution vient souvent d'une critique externe à l'institution, voire au sein de l'institution même.
- 55 Sur la question du chiffre je suis d'accord. On peut faire des choses avec les données numériques. Je sais qu'il y a beaucoup d'outils extrêmement efficaces mais le risque c'est de s'en contenter. Ce que je mets en avant c'est le fantasme des big data.
- 56 Jean-Luc Pinol : Là on est d'accord. Mais simplement, je pensais au dernier ouvrage de Franco Moretti sur l'histoire littéraire où il dit, de façon très provocatrice, arrêtez de lire, comptez ! Son livre, c'est en français *Graphes, cartes et arbres*⁶. Il montre bien, d'ailleurs avec des méthodes que pour autant je peux considérer, en termes de cartographie, assez frustrées, comment on est obligés de repenser la manière de faire de la recherche.
- 57 Marjorie Burghart : Je voulais revenir sur le côté *steampunk*, qui m'a beaucoup plu. Sur une de tes premières diapositives, tu citais ce passage sur les *Current scholarly publishing*, et je me demandais où était le « steam » et où était le « punk » dans tout cela. Quand on entend cela, on imagine des gens qui amènent leurs trucs tout poussiéreux et qui les

donnent à des gens très pointus que l'on voit presque avec des costumes de la NASA en train de se demander ce qu'ils vont en faire Est-ce que l'on n'est pas en fait dans la situation tout à fait inverse, une situation où des personnes sont en train de faire leur révolution industrielle, des gens tout à fait à la pointe au niveau intellectuel, et se confrontent à quelque chose de très vapeur, à quelque chose de très victorien qui rend les projets très lourds à mener. Voilà, il y a un anachronisme c'est sûr. Mais, je me demandais de quel côté sont les humanités, et de quel côté est le digital dans cet anachronisme.

- 58 Paul Bertrand : Je suis tout à fait d'accord. On peut poser la question dans les deux sens. La façon dont je l'avais entrevue, comme Andrew Reinhard l'avait entrevue, c'était dans ce premier sens. Mais on pourrait tout à fait envisager les choses dans ce sens-là aussi. Le *steampunk* est dans les deux sens, tout à fait.
- 59 Paul Girard : Je vous remercie pour cette intervention. Je vais plutôt donner le point de vue d'un ingénieur, du coup. J'ai entendu beaucoup de questions et de problématiques mais en fait peu de réponses concrètes sur la question des données. Ne serait-ce pas justement une bonne pratique que de définir ce qui devrait être la « bonne donnée » pour les *digital humanities* ? Il y a en tout cas probablement un juste milieu entre le quanti et le quali qui est celui que vous proposez à la fin sous la forme de laboratoires comme ceux qui existent à Sciences-Po ou à l'ENS Lyon. Il y a probablement un juste milieu entre le PDF et la TEL. On peut avoir un texte qui ne soit pas un texte brut et qui pour autant soit traitable par des informaticiens, et qui ne soit pas bloqué comme le PDF qui malheureusement est ingérable pour les ingénieurs. N'y a-t-il pas à définir, à défaut de bonnes pratiques, de bons principes pour l'interopérabilité entre les chercheurs ?
- 60 Paul Bertrand : Je pense qu'on peut même aller encore plus loin. Je pense, que ce qu'il faudrait vraiment – au-delà des formats, au-delà des problèmes que j'ai soulevé – c'est se poser la question de savoir ce que c'est que ce donné que nous avons. Avant même d'essayer de le formater, ou de le mettre dans un format, je voudrais simplement que l'on réfléchisse. C'est quelque chose que tout le monde attend et qui permettrait à la communauté des *digital humanities* de savoir jusqu'où aller plus loin, et peut-être de choisir ces bonnes voies. Mais avant de choisir ces bonnes voies, il faut déjà savoir sur quoi on travaille. Et pour l'instant, moi, cela fait des années que je ne vois pas cette réflexion se mettre en place. Non parce qu'il n'y a pas de volonté mais parce que l'on n'a pas le temps de le faire, parce que l'on a une pression énorme qui nous empêche de mener cette réflexion. Ce que je propose c'est de faire un non-laboratoire en quelque sorte ; un peu dans l'idée du THATCamp, l'idée d'un espace hors du temps, hors du temps de l'institution, même si c'est un fantasme, dans lequel on tente de réfléchir à ce donné.
- 61 Stéphane Pouyllau : Je suis principalement d'accord avec tout ce que tu as dit tout au long de ton exposé, cela n'est pas vraiment une surprise. Sur cette problématique autour de la donnée, je suis assez d'accord avec l'idée du non-laboratoire, l'idée d'avoir un espace de réflexion et de discussion qui se réunisse à intervalle régulier. Le problème, c'est que dès qu'on va le réunir à intervalle régulier, on va l'institutionnaliser. On sera toujours, je pense, obligés de passer par ce genre de choses. Mais je pense que c'est essentiel parce que la question des données et de la capacité qu'ont les chercheurs à les structurer, à les malaxer pour les réutiliser, est liée à un besoin qu'il faut qu'on arrive à faire émerger, et à carrosser sur le plan de la formation pour qu'on ait à tous points du territoire, dans tous les laboratoires, à un moment donné quelqu'un qui dise « attention, je dois me préoccuper de la réflexion sur les données ». Et là où je te rejoins, c'est que la pression institutionnelle l'empêche et agit comme un couvercle qui tombe par-dessus ces

réflexions. Il est très important de relever ce couvercle, pour donner de l'air et de retrouver du temps pour savoir ce que l'on est en train de faire avec ces données et comment on va faire avec les données que l'on a à notre disposition. Cela me semble essentiel, et je pense qu'il faut arriver à trouver une solution ; je souscris à cette idée de non-laboratoire, ou de laboratoire virtuel qui permettrait de prendre ce temps-là.

- 62 Par ailleurs, je pense qu'il y a quelque chose à inventer pour que les collègues qui utilisent les bases de données, quel que soit le format (je ne me positionne pas en termes de formats ou de logiciels utilisés) arrivent à se mettre dans la situation où ils sont toujours dans une position de chef de projet, en assistance à la maîtrise d'ouvrage. Il faut qu'ils soient capables de mobiliser rapidement leurs compétences pour produire une base de données mais qu'ils aient conscience qu'à un moment donné il faut qu'ils fassent appel à d'autres personnes pour faire le travail parce que celui-ci sort de leurs compétences. C'est la notion de relation : savoir à partir de quel moment on va devoir passer le témoin à des gens qui ne sont pas issus du monde académique, qui ne connaissent pas le terrain, mais que l'on va aller chercher parce qu'ils savent faire, ou qu'ils savent faire autre chose que l'on ne sait pas faire. Et de la tension qui va s'installer, de la discussion qui va s'installer entre eux et le chercheur, c'est de là que va naître en fait l'innovation. Il y a un travail à faire auprès des collègues historiens, géographes, ou du monde académique en sciences humaines et sociales, pour qu'il se comportent en chefs de projets. Qu'ils sachent faire une partie du travail, très bien, mais à un moment donné aient conscience qu'il faut chercher d'autres compétences ailleurs.
- 63 Paul Bertrand : Je ne peux qu'être d'accord avec toi, c'est évident. Pour le non-espace, ce que je veux vraiment dire c'est que l'on relance une démarche d'innovation – je sais bien que le terme est très connoté, très politiquement correct et en même temps pas très intéressant. Mais derrière le concept d'innovation, il y a la recherche. La recherche qui, par nature, tend toujours à aller plus loin. L'idée c'est de relancer une démarche un peu audacieuse parce que je pense que l'on en manque assez ces derniers temps.
- 64 On est souvent confronté au problème très concret du contact entre les métiers. Faire en sorte que le chercheur comprenne qu'un ingénieur *digital humanities* n'est pas là pour publier son corpus, mais qu'il est là pour accompagner son corpus du début à la fin. L'idée c'est d'accompagner. Et quand les problèmes d'accompagnement seront passés dans les mœurs alors la question sera résolue ; presque.
- 65 Frédéric Clavert : Ma question a été un peu vidée de son sens par celle de Stéphane. Mais peut-être une simple question, concrètement, ce non-laboratoire, qu'est-ce que c'est exactement ? Comment l'organise-t-on, enfin on ne l'organise pas, c'est un non-laboratoire : mais qu'est-ce que l'on fait ?
- 66 Paul Bertrand : Sur comment l'on fait, je n'ai aucune idée. Sur ce qu'on y ferait, là j'ai une petite idée. L'idée que j'ai c'est de se retrouver d'abord simplement autour de nos données ; sur ce que nous entendons chacun par « données ». Puis d'aller plus loin, en mettant à plat, chacun en essayant de décortiquer la réalité face à laquelle nous nous trouvons ; tout en essayant de le raccrocher avec les projets dans lesquels nous sommes, essayer de trouver des points de convergences, des points de ralliement. C'est une démarche très philosophique, très intellectualisante que je propose là. Mais vous savez bien que les *digital humanities* c'est une approche très philosophique de la donnée, ce n'est pas beaucoup autre chose. Il faudrait donc travailler sur les données que nous rencontrons et essayer de les classer, de les répertorier, non pas à la façon de Linée, mais

dans une perspective de structuration sémantique, matérielle éventuellement, documentaire, etc. C'est à décliner, je lance les choses, mais tout est à construire.

67 Frédéric Clavert : Un atelier demain là-dessus ?

68 Paul Bertrand : Pourquoi pas. C'est un peu l'idée que j'avais en tête. Mais j'attendais les tomates ou pas.



69 Pierre Mounier : Je suis content de voir que cette idée d'une démarche beaucoup plus réflexive recoupe beaucoup les textes qui ont été choisis pour le *Read / Write Book 2*. Vous verrez, beaucoup de textes ont été choisis pour cette dimension là, à la fois réflexive et critique sur ce que deviennent, ou sont en train de devenir, sous un certain nombre de pressions, les *digital humanities*. Je pense que c'est une démarche vraiment importante et je suis ravi qu'on la partage.

70 Deuxième chose, je peux annoncer que si les acteurs des *digital humanities* ont découvert le *steampunk* grâce à Boulet, il faut savoir aussi qu'il y a une demi-heure, Boulet a découvert les *digital humanities* sur twitter, en particulier grâce à Rémi Mathis qui l'a tweeté, et Boulet a répondu !

71 Ce média ne peut être affiché ici. Veuillez vous reporter à l'édition en ligne <http://books.openedition.org/editionsmsh/327>

NOTES

1. <http://classicslibrarian.wordpress.com/2012/06/04/lawdi-conference-on-linked-open-data-for-ancient-studies/>?
 2. *Read/Write Book 2, Une introduction aux humanités numériques*, Pierre Mounier (dir.), Openedition press, 2012, <http://press.openedition.org/226>, page consultée le 27 septembre 2012.
 3. <http://www.rechercheisidore.fr>
 4. Harmut Rosa, *Accélération, Une critique sociale du temps*, coll. Théorie Critique, Paris : Éditions La Découverte, 2010.
 5. Langlois et Seignobos, *Introduction aux études historiques*, 1898.
 6. Franco Moretti, *Graphes, cartes et arbres, Modèles abstraits pour une autre histoire de la littérature*, Paris : Éditions Les prairies ordinaires, 2008.
-

INDEX

Mots-clés : steampunk, digital humanities, cyberpunk, recherche, donnée, métadonnée